Safety Science 102 (2018) 211–225

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Addressing the epistemic uncertainty in maritime accidents modelling using Bayesian network with interval probabilities

Guizhen Zhang[a], Vinh V. Thai[b,*], Kum Fai Yuen[c], Hui Shan Loh[d], Qingji Zhou[e]

[a] DHI-NTU, Nanyang Environment and Water Research Institute, Interdisciplinary Graduate School, Nanyang Technological University, 1 Clean Tech Loop, Singapore 637141, Singapore
[b] School of Business IT & Logistics, RMIT University, Melbourne, VIC 3000, Australia
[c] Department of International Logistics, College of Business & Economics, Chung-Ang University, 84 Heukseok-ro, Heukseok-dong, Dongjak-gu, Seoul 156756, South Korea
[d] Logistics and Supply Chain Management Programme, School of Business, Singapore University of Social Sciences, 461 Clementi Road, Block C, Singapore 599491, Singapore
[e] School of Civil and Environmental Engineering, Nanyang Technological University, Blk N1.1-B3-01, 50 Nanyang Avenue, Singapore 639798, Singapore

ARTICLE INFO

Keywords:
Interval probabilities
Bayesian network
Maritime accidents
Experts' elicitation
Epistemic uncertainty

ABSTRACT

Bayesian Network (BN) is often criticized for demanding a large number of crisp/exact/precise conditional probability numbers which, due to the lack of statistics, have to be obtained through experts' judgment. These exact probability numbers provided by the experts often carry a high level of epistemic uncertainty due to the incompleteness of human knowledge, not to mention the hardness in obtaining them in the first place. The existence of uncertainty in risk modelling was well recognized but seldom discussed. This paper explores the extension of BN with interval probabilities to the modelling of maritime accidents, which allows for the quantification of the epistemic uncertainty. Ship collision is chosen for case study for the strategic importance of navigational safety. The user friendly linguistic terms defined with interval scales were used for elicitation of interval conditional probabilities from industry experts. Inferences were made directly with the interval probabilities with the GL2U algorithm. Meanwhile, the interval probabilities were converted into point probabilities and computed with the traditional BN method for comparison, which were all shown to be within the ranges of the calculated posterior intervals probability. Results with inputs from different experts reveal discrepancies, which in turn verify the existence of uncertainty in risk modelling. A discussion was also provided on how the uncertainty in risk assessment propagates to the decision making process and influences the ranking of potential risk control options.

1. Introduction

1.1. Maritime accidents and BN

Maritime accidents have continued to occur, which threaten the safety of seafarers at sea, the economic performance of shipping companies and the environment. Therefore, understanding why and how accidents happen is of great importance for future safety management. Since accidents cannot be completely avoided, the reasonable goal is to control the accident risk to a desired level. Risk assessment is essential for this purpose. By performing risk analysis, we can evaluate the safety level of the current system as well as identifying the most critical issues. Some risk assessment methods also enable the evaluation of risk control options and thus ascertaining the most cost-effective way for reducing the risk level. Many risk analysis methods have been developed in the past few years, including Hazard and Operability Studies (HAZOP), Failure Mode and Effects Analysis (FMEA), Event Tree Analysis (ETA), Fault Tree Analysis (FTA) and Bayesian Belief Network (BN). Each of these risk analysis tools has its unique characteristics and fits different purposes.

BN is becoming an increasingly popular methodology for risk analysis of the maritime transportation system in recent years due to its capability to model causal interdependence, to incorporate of experts'

Abbreviations: AHP, Analytic Hierarchy Process; AIS, Automatic Identification System; ATSB, Australian Transport Safety Bureau; BN, Bayesian Network; BRM, Bridge Resource Management; DNV, Det Norske Veritas; ETA, Event Tree Analysis; FMEA, Failure Mode Effects Analysis; FSA, Formal Safety Assessment; FTA, Fault Tree Analysis; GCAF, Gross Cost of Averting a Fatality; GL2U, Generalized Loopy 2-Updating; HAZOP, Hazard and Operability Analysis; HFACS, Human Factors Analysis and Classification System; IMO,, International Maritime Organization; IPE, Iterated Partial Evaluation; IPT, Interval Probability Theory; MAIB, Marine Accident Investigation Branch; NCAF, Net Cost of Averting a Fatality; OOW, Officer on Watch; RCO, Risk Control Options; SLP, Support Logic Programming; STRAITREP, The Mandatory Ship Reporting System in the Straits of Malacca and Singapore; SV2U, Structured Variational 2U Method; TSB, Transportation Safety Board of Canada; VTS, Vessel Traffic Service

* Corresponding author.
E-mail addresses: GZHANG004@e.ntu.edu.sg (G. Zhang), vinh.thai@rmit.edu.au (V.V. Thai), yuenkf@cau.ac.kr (K.F. Yuen), hsloh@suss.edu.sg (H.S. Loh), qjzhou@ntu.edu.sg (Q. Zhou).

http://dx.doi.org/10.1016/j.ssci.2017.10.016
Received 29 September 2016; Received in revised form 24 August 2017; Accepted 18 October 2017
Available online 04 November 2017
0925-7535/

Table 1
Comparison of aleatory uncertainty and epistemic uncertainty.

Comparison	Aleatory uncertainty	Epistemic uncertainty
Source of the uncertainty	Variability of the underlying stochastic process	Incomplete knowledge of the system
Existence	Quantities	Model and parameter
Associated with	System	Analyst
Reducible	Not reducible	Reducible
Modelling by	Probability number/distribution	Alternative probability number/distribution
Equivalent terms	Variability, natural uncertainty, objective uncertainty, inherent variability, (basic) randomness, and type-A uncertainty	Subjective uncertainty, lack-of-knowledge or limited-knowledge uncertainty, ignorance, specification error, prediction error, and type-B uncertainty

knowledge when statistical data does not exist, to make dynamic updates when new observation is made, and to include human and organizational factors. Helle et al. (2011), Lehikoinen et al. (2015), Montewka et al. (2014), Banda et al. (2016) and Zhang et al. (2013) are a few examples of BN applications in the maritime risk analysis field. A more detailed review of the literature on maritime accidents risk prediction based on Bayesian Network can be referred to Goerlandt and Montewka (2015b) as well as Zhang and Thai (2016). BN was also recommended for risk assessment (Step 3 of the Formal Safety Assessment, FSA) to International Maritime Organization (IMO) (IMO, 2006). The application of BN includes three steps, i.e. BN structure development, parameterization, and inferences. Both the BN structure and parameters could be built manually, automatically or a combination of both (Neil et al., 2000; Kjrćulff and Madsen, 2013; Neil et al., 2000).

1.2. Uncertainty and BN modelling

The consideration of uncertainties is crucial for obtaining reliable results in risk analysis (Merz and Thieken, 2005). By sources, uncertainties could be broadly separated into two class: aleatory and epistemic uncertainty, the comparison of which could be found in Table 1.

Liu et al. (2003) reviewed some of the most important uncertainty reasoning approaches, including the Bayesian theory of probability, Dempster-Shafer theory of evidence, and fuzzy set theory. Each of these approaches views and handles uncertainties from different perspectives. Bayesian Theory has many good features such as strong theoretical root, less computational complexity compared with other approaches. It models aleatory uncertainty through probability but could not include epistemic uncertainty since each entity must be assigned with exact probability numbers.

Due to the lack of available statistical data, experts' opinion is an important source for the probability specification or parameterization in BN modelling, especially for applications to the maritime risk assessments. This, however, poses huge challenges for the reliability of the model as well as the involved domain experts. First, for probability elicitation, the experts are often asked about the conditional dependence between the model elements on top of their own expertise. Moreover, the requirement to elicit a large number of probability numbers adds to the workload of the experts. From the viewpoint of modelling, the involvement of experts will lead to epistemic uncertainty, due to the lack of knowledge about the system (Liu et al., 2003; Merrick et al., 2005; Fallet et al., 2011), which can sometimes be referred to as quantities which have fix values, but their exact value are unknown (Swiler et al., 2009).

The lack of systematic consideration of uncertainty in the applications of maritime transportation risk analysis was identified through a detailed review in Goerlandt and Montewka (2015b), even though the existence of uncertainties are recognized and accepted. One exception was Merrick et al. (2005) which used Bayesian approach to estimate the impacts of parameter uncertainties in the traffic simulation model (not the risk analysis model) for evaluating the ferry expansion alternatives. In the last few years, there have been more studies with focus on uncertainties in the maritime risk models. For example, Sormunen et al. (2014) showed through an extensive study that the uncertainties in accident and risk models can be significant. Goerlandt and Montewka (2015a) went one step further by introducing a framework where uncertainties are qualitatively assessed. However, so far, there are no studies on maritime risk analysis which quantitatively address the epistemic uncertainties. The objective of this paper is to provide a method to model the epistemic uncertainties related with the probability parameters in Bayesian Network models for maritime risk analysis.

To achieve the objective, this paper seeks to extend BN by including interval probabilities. Interval probability expresses imprecision in a more straightforward way. Fallet et al. (2011) concluded that the interval probability method best represents the experts' knowledge due to more appropriate semantics as compared to hard evidence, soft evidence, and total ignorance. In applications, obtaining interval probability parameters are much easier than getting point probabilities, especially when there is only little, incomplete or conflicting information available to assist experts' judgment (Guo and Tanaka, 2010). In other cases, when multiple experts are involved, each expert may indicate their own belief and if no consensus could be reached, the result are interval probabilities (Cozman, 2000). Considering these facts, the application of interval probability (with an upper bound and lower bound) in BN could bring more application value.

1.3. Organization of this paper

The rest of this paper is organized as follows. Section 2 defines interval probability, discusses its properties and summarizes the updating algorithm for BN with interval probabilities. Section 3 presents the application of BN with interval probabilities to ship collision causation probability modelling. The detailed elicitation process is also discussed in this section. Section 4 shows the inference result with the interval probabilities. An example of the influence on the evaluation of risk control options with interval probabilities is provided as well. Finally, Section 5 summarizes the paper.

2. Methodology

This paper extends the traditional BN to include interval probability parameters for maritime risk modelling. Inferences are made directly with interval parameters. The following subsections present the definition of credal network, interval probability and the relative properties.

2.1. Credal network and BN with interval probability parameters

BN with interval probabilities is a special type of credal network, which extends BN to deal with imprecision and uncertainty (Corani et al., 2012). A credal network over a set of random variables $\mathbf{X} = (X_1, \ldots, X_k)$ is $\langle G, \{\mathbf{P_1}, \ldots, \mathbf{P_m}\}\rangle$, where G is a directed acyclic graph whose nodes have one-to-one correspondence to the elements in $\mathbf{X}$, and $\langle G, \mathbf{P}_j\rangle$ is a BN over $\mathbf{X}$ for each $j = 1, \ldots, m$ (Antonucci, 2008; Antonucci and Zaffalon, 2008). This definition indicates that a credal network could be regarded as a set of BNs, as illustrated in Fig. 1.

Fig. 1a is an example of the traditional BN where all the probabilities are exact numbers. Fig. 1b is a credal network with the same structure. The probabilities in the credal network are imprecise, being an interval or comparisons of the probabilities, which enables the representation of classificatory and comparative probability judgements

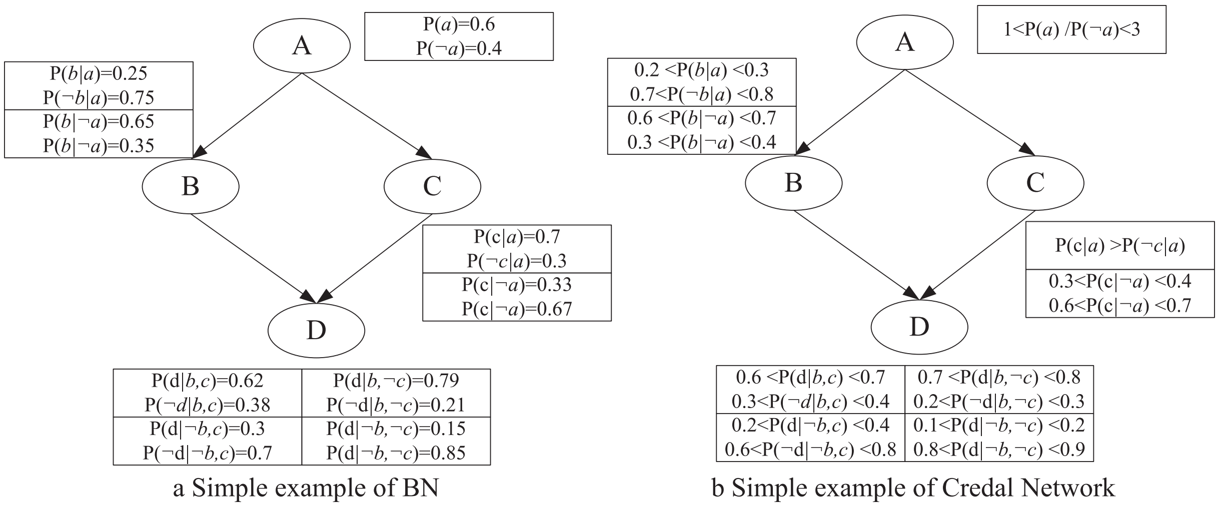

Fig. 1. (a) Simple example of BN; (b) simple example of Credal Network.

(Piatti et al., 2010). For example, for node A, the judgement $1 < P(a)/P(\neg a) < 3$ means that the chance of A = a is one to three times higher than the chance A = $\neg a$. For node B, the probability of P(b|a) could be any value between 0.2 and 0.3. The BN in Fig. 1a is just one among many others that satisfy the probability conditions of the credal network in Fig. 1b. The inference with a credal network is the same with inferences with its vertices (Antonucci, 2008). However, the number of vertices is exponential to the input size except for the case of binary nodes. The complexity in inference for credal network is thus greatly increased than traditional BN.

2.2. Definition and properties of interval probability

For a random variable X which takes values from a finite set $\{x_1, \ldots, x_n\}$, the intervals $L = \{L_i = [L(a_i), U(a_i)], i = 1, 2, \ldots, n\}$ are called the interval probability (Guo and Tanaka, 2010; Hu et al., 2012) if and only if for any $P(a_i) \in L_i$, there exists $P(a_j) \in L_j$, such that,

$$P(a_i) + \sum_{j=1,2,\cdots i-1,i+1,\cdots n} P(a_j) = 1. \tag{1}$$

L will satisfy (1) if and only if they satisfy (2) and (3), where $i, j \in [1, \ldots, n]$) (Tessem, 1992; de Campos et al., 1994; Weichselberger, 2000; Guo and Tanaka, 2010; Tessem, 1992):

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} L(a_i) + U(a_j) \leqslant 1 \tag{2}$$

$$\sum_{\substack{i=1 \\ i \neq j}}^{n} U(a_i) + L(a_j) \geqslant 1 \tag{3}$$

Normally, the elicited interval probabilities may or may not satisfy (2) and (3). However, it is easy to check whether they satisfy condition (4):

$$\sum_{i=1}^{n} L(a_i) \leqslant 1 \leqslant \sum_{i=1}^{n} U(a_i) \tag{4}$$

Condition (4) is a necessary but insufficient condition of (2) and (3). The intervals that satisfies condition (4) are called semi-interval probabilities, denoted with $[L'(a_i), U'(a_i)]$. Interval probabilities could be elicited from $[L'(a_i), U'(a_i)]$ by solving the linear programming problem as shown in function (5) (Guo and Tanaka, 2010):

$$Max \sum_{i=1,2,\ldots n}^{n} (U(a_i)-L(a_i)) s.t. \sum_{\substack{i=1 \\ i \neq j}}^{n} L(a_i) + U(a_j) \leqslant 1, \sum_{\substack{i=1 \\ i \neq j}}^{n} U(a_i) + L(a_j) \geqslant 1 U(a_i) \geqslant L(a_i), U(a_i) \leqslant U'(a_i), L(a_i) \geqslant L'(a_i) \tag{5}$$

In the case of binary variables ($n = 2$). When $j = 1$, we can obtain (6) and (7):

$$L(a_2) + U(a_1) \leqslant 1 \tag{6}$$

$$U(a_2) + L(a_1) \geqslant 1 \tag{7}$$

Similarly, when $j = 2$, we can obtain (8) and (9):

$$L(a_1) + U(a_2) \leqslant 1 \tag{8}$$

$$U(a_1) + L(a_2) \geqslant 1 \tag{9}$$

Combining (6) and (9), $L(a_2) + U(a_1) = 1$; similarly, from (7) and (8), $L(a_1) + U(a_2) = 1$ (Hall et al., 2005), which implies that the boundary for one state could be inferred from the boundary of the other. For example, if L_1 is $[L(a_1), U(a_1)] = [0.3, 0.6]$ for a_1, then L_2 can be calculated as $[L(a_2), U(a_2)] = [1–0.6, 1–0.3] = [0.4, 0.7]$. The calculated interval probability not only speeds up the elicitation process but also ensures consistency.

2.3. Combination of experts' opinion

When multiple experts are consulted for probability elicitation, either group consensus or individual opinions may be obtained as the output. The individual judgement can be combined in various ways, from simply adopting a large interval to using the simple or weighted average. The weighted average method was adopted in this study. Here, a weight number is assigned to each judgment, i.e. each interval probability, rather than to each expert. Therefore, the weight for the same expert may vary for the different judgments they made. It is assumed that more weights should be assigned when one judgment is closer to other judgments. More specifically, the weight of the interval probability specified by one expert is reciprocally proportional to the distance between the interval to all other intervals specified by other experts. The distance L^b between interval $[L^b(a_i), U^b(a_i)]$ and all other intervals $[L^j(a_i), U^j(a_i)]$, where $j = 1, 2 \ldots b-1, b+1, \ldots m$ (m is the total number of experts) is calculated as follows (Hu et al., 2012):

$$L^b = \frac{1}{m-1}\sum_{\substack{j=1\\ j\neq b}}^{m}\sqrt{(L^j(a_i)-L^b(a_i))^2 + (U^j(a_i)-U^b(a_i))^2} \tag{10}$$

As such, the weight for expert b was defined as k/L^b, where k is a constant number. All the weights should add up to 1, i.e.:

$$\sum_{b=1,2,\dots m} k/L^b = 1 \tag{11}$$

From Eq. (11) value of k could be calculated. The weight assigned in this method is objective. The combined interval probability is closer to the majority of judgments.

2.4. Updating BN with interval probabilities

Updating credal networks is much more complicated than updating the traditional BN. For simplicity, Hu et al. (2012) converted the interval probabilities to exact probability number and then apply traditional BN method for analysis. Similarly, Ge et al. (2013) decomposed the interval-valued BN into two traditional point-valued BNs, one with the upper bound and one with the lower bound values, using the BN software GeNie (Bayesian Fusion, 2017) to make inferences. Meanwhile, many algorithms have been proposed for making inferences directly using interval probability in belief networks. Hall et al. (2005) used Support Logic Programming (SLP) and Interval Probability Theory (IPT) for analysis of interval probabilities in BN in the study about flood and climate change. Liu and Yue (2011) extended the Gibbs sampling algorithm to the updating of interval probabilities. Antonucci et al. (2013a) proposed a linear programming method for approximate updating of credal networks.

The 2U algorithm extended from Pearl's BN updating algorithm is a widely used algorithm (Fagiuoli and Zaffalon, 1998). It is an exact algorithm for updating of binary polytree credal networks. Ide and Cozman (2004) extended the 2U method to multi-connected networks, i.e. Loopy 2U. Even though 2U is an exact algorithm, L2U is only an approximate algorithm. L2U performs best compared to the other approximate inference methods: IPE (Iterated Partial Evaluation) and SV2U (Structured Variational 2U) in terms of time and accuracy (Ide and Cozman, 2008). GL2U further develops L2U to deal with nonbinary credal network through the binarization algorithms which transfers any credal network into an equivalent binary credal network (Antonucci, 2008; Antonucci et al., 2010). The only approximation in GL2U is related to the loopy step. In this paper, the GL2U algorithm is applied for calculation and analysis. More details of the GL2U algorithm could be found in Antonucci (2008) and Antonucci et al. (2010).

3. Application of Bayesian network with interval probabilities to maritime accidents modelling

3.1. Reviews of risk assessment models for ship collision accidents

Ship collision is chosen for case study for its strategic importance on navigational safety (Kuehmayer, 2008; Kujala et al., 2009). For collision risk modelling, analysts adopt quite different perspectives and approaches. There are quantitative as well as qualitative methods. Some authors consider probability and consequence separately while some adopt integrated approaches. Risk indicator, probability numbers are among some of the measurements. Readers can be directed to Goerlandt and Montewka (2015b) for a more detailed discussion. In this paper, one of the most widely used probabilistic risk approach is used for the case study, where collision probability is calculated as the product of the Geometric probability and causation probability. The most common approach is to calculate the collision probability as the product of the Geometric probability and causation probability (Li et al., 2012; Kuehmayer, 2008).

$$P = P_a \times P_c \tag{12}$$

Geometric probability refers to the probability of a ship becoming a collision candidate, for whom a collision would happen if no further evasive actions are taken. Causation probability is the probability that a collision candidate fails to conduct any aversive actions and thus results in grounding or collision. Geometric probability and causation probability are normally calculated separately and then multiplied together. Geometric probability is calculated either adopting analytical models (Fujii et al., 1974; Pedersen, 2010) or using dynamic models with system simulation of the real traffic (Goerlandt and Kujala, 2011). Meanwhile, causation probability is computed with accident statistics, adjusting previous values elicited from experts' opinion, modelled with Fault Tree Analysis and BN (Mazaheri, 2009; Fujii et al., 1974), or using dynamic models with system simulation of the real traffic.

To estimate collision frequency for a specific geographical location, detailed geometric models are needed while assumed values of causation probabilities are adopted. On the other hand, for understanding accident causal relations, detailed causation probability model is essential. The focus of this research is on the latter, i.e. modelling of causation probability.

3.2. BN model structure for ship collision causation probability

Detailed collision causal narratives can be found from accident investigation reports by marine accident investigation authorities such as Marine Accident Investigation Branch UK (MAIB), Transportation Safety Board of Canada (TSB) and Australian Transport Safety Bureau (ATSB), which records, rebuilds and analyzes each accident in details. For BN structure construction, the accident causal chain (Grabowski, 2000), Reason's Swiss cheese model and the Human Factors Analysis and Classification System (HFACS) have been used as theoretical frameworks (Ren et al., 2008, 2009; Akhtar and Utne, 2014). The BN model structure in this research was built after reviewing many accident reports. References to previous models in the literature were also made, especially the model by Friis-Hansen and Simonsen (2002) and Det Norske Veritas (2003).

Under a critical encounter situation, the three important cognitive functions for successful collision avoidance identified by Leva et al. (2006) are: Detection-Interpretation, Interpretation-Planning and Execution of the actual maneuvering. Similarly, Montewka et al. (2017) decomposed the evasive action into three components, i.e. detection, assessment and action. These concepts were adopted in the present model. Three nodes were incorporated, including "Detection", "Maneuver planning" and "Maneuvering", corresponding to the three cognitive functions. Successful "Detection" could be achieved through "Navigation system detection", "Visual detection" or "Vessel Traffic Service (VTS) detection". Meanwhile, successful detection is the prerequisite for correct "Maneuver planning", which in turn is the prerequisite for correct "Maneuvering" action. Collision is determined by the combined impact of the "Maneuvering" action from both encountering ships. The logic is reflected in the conditional probabilities, as in Table 2.

The various causal factors such as vessel specifications, route characteristics, weather conditions, human and organizational factors were incorporated into the model by exerting influence on the critical functions (Mazaheri, 2009). Since the two-ship collision scenario is modelled in this research, two sets of variables with the same definition and same causal structure corresponding to each ship were used in the model. The total number of nodes was controlled to a manageable size considering the elicitation workload. The detailed model is presented in Fig. 2.

Table 2
Conditional probability for the node "Detection"

Navigation system detection	Yes				No			
Visual detection	Yes		No		Yes		No	
VTS detection	Yes	No	Yes	No	Yes	No	Yes	No
Detection = Yes	1	1	1	1	1	1	1	0
Detection = No	0	0	0	0	0	0	0	1

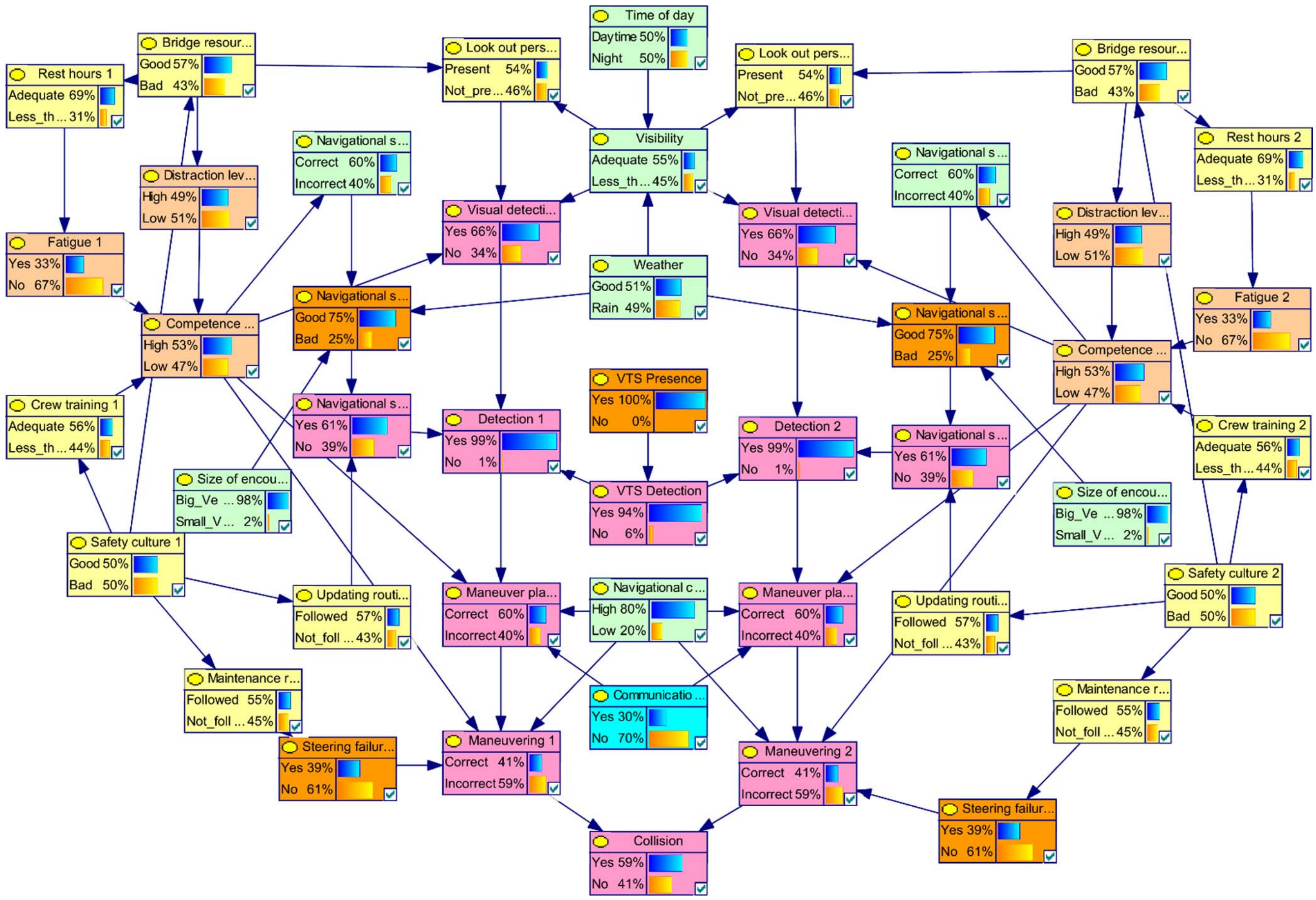

Fig. 2. Collision model based on causal relations.

The eight nodes i.e. "Time of day", "Visibility", "Weather", "VTS Presence", "VTS Detection", "Navigational complexity", "Communication between ships" and "Collision" are the common variables influencing both ships.

The states and a brief description for all nodes in the model are presented in Table 3.

3.3. The elicitation of conditional probabilities from experts

Due to the lack of statistical data, experts' elicitation plays an important role in obtaining the conditional probability numbers. A detailed literature review on probability elicitation in Bayesian Network modelling of maritime accidents was conducted in a recent study (Zhang and Thai, 2016). In this research, to alleviate the elicitation workload, several techniques were employed. First, for the two sets of similar variables, elicitation was only needed for one set as the result of which applies to the other set. Second, since all nodes in the collision model are binary nodes, elicitation is performed for one state only. The interval probability for the other state could be computed using $L(a_2) + U(a_1) = 1$ and $L(a_1) + U(a_2) = 1$, as discussed in Section 2.1. Third, for the node "Detection", the probability was specified according to logic.

For individual probability elicitation, either direct or indirect methods can be used (Kuhnert et al., 2010), such as the probability wheel, the probability scale (Renooij, 2001), the gambling analogy, Analytical Hierarchy Process (AHP) (Wang et al., 2010) and the fuzzy methods (Ren et al., 2008, 2009). Among them, linguistic terms such as "Likely", "Very likely", "Virtually certain", "Medium likelihood", "Unlikely", "Very unlikely" and "Extremely unlikely" associated with probability interval were found to be user-friendly (Mastrandrea et al., 2010; Antonucci et al., 2013b; Ren et al., 2007). The natural language corresponds to human cognition, and can represent imprecise/vague information. The scaling of linguistic terms from three literature sources were compared in Table 4 and the definition from the first source was adopted in this paper (Mastrandrea et al., 2010).

A questionnaire was developed in this research for elicitation, an excerpt of which is shown in Fig. 3. The definition of linguistic terms with the corresponding probability intervals were explained both verbally and in a figure to experts. The experts could choose the likelihood from a dropdown list, which is comparatively easy and time-saving.

In total, eleven completed questionnaires were collected after face-to-face interviews with navigational experts in Singapore. The average time of the elicitation process ranges from one to two hours. Assumptions were made for probability distributions for nodes without parent nodes, as demonstrated in Table 5. The Singapore Strait was used as the background for the study. The prior probabilities here are point probabilities although they could also take the form of interval probabilities.

4. Results and discussions

4.1. Marginal interval probabilities

Inferences were made directly with interval conditional probabilities using the GL2U algorithm. First, the calculation was performed using the combined interval conditional probabilities obtained from all experts. For comparison, the combined interval conditional probabilities were converted to point probabilities and used as input for computations with the traditional BN method. The results are summarized in Table 6. Column 2 and 5 show the result of the marginal probabilities computed with the traditional BN while Column 3 and 6 list the marginal interval probabilities computed with the GL2U method. It could be observed that all the point probabilities computed with traditional Bayesian network method fall into the probability intervals obtained from the GL2U methods.

Fig. 4 shows the model with the interval marginal probabilities, where

Table 3
Causal factors of the ship collision model.

Node	States	Description
Collision	Yes; No	Whether a collision happens or not
Maneuvering 1(2)	Correct; Incorrect	Whether the ship performed the desired avoidance action according to COLREG
Maneuver planning 1(2)	Correct; Incorrect	Whether the ship makes a proper emergency passing plan after detecting the other ship
Updating routine 1(2)	Followed; Not followed	Whether the OOW frequently checks and updates of the navigational equipment
Navigational system signal 1(2)	Good; Bad	Whether there is strong/stable signal on the navigational system
Navigational System settings 1(2)	Correct; Incorrect	Whether Navigational system such as the Radar and Automatic Identification System AIS were set to the right range and tuned properly
Navigational system detection 1(2)	Yes; No	Whether the other ship was detected by the OOW through the navigation system
Look out person 1(2)	Yes; No	Whether an additional look out person is assigned
Visual detection 1(2)	Yes; No	Whether the other ship was detected by the OOW visually
VTS Presence	Yes; No	Whether Vessel Traffic Service is provided in the considered area
VTS Detection	Yes; No	Whether the potential collision is detected by VTS and communicated to OOW
Detection 1(2)	Yes; No	Whether the other ship was detected through the navigation system, visually or by VTS and communicated to OOW
Rest hours 1(2)	Adequate; Less than adequate	Whether the OOW got adequate rest. According to STCW convention, seafarers should get a minimum of 10 hours' rest in any 24-h period and 77 h in any 7-day period
Fatigue 1(2)	Yes; No	Whether the OOW is fatigued. Fatigue will reduce the competence of the OOW
Distraction level 1(2)	High; Low	Level of distraction of OOW by administrative tasks, talking with people on board or over the phone, etc.
Crew training 1(2)	Adequate; Less than adequate	Whether the company identified the training needs of their crew and provide the relevant training
Competence of the OOW 1(2)	High; Low	Refers to the knowledge and skills as well as the physical and mental state of the OOW
Safety culture 1(2)	Good; Bad	Safety culture is the attitude, beliefs, perceptions and values that employees share in relation to safety in the organization
Maintenance routine 1(2)	Followed; Not followed	Whether the ship carries out maintenance in intervals according to manufacturers and class/Classification society requirements. Shipping companies should have planned maintenance system on ships as per ISM (International Safety Management Code)
Bridge Resource Management BRM 1(2)	Good; Bad	BRM is the effective management and utilization of all resources, human and technical, available to the bridge team, to ensure the safe completion of the vessel's voyage bridge design
Weather	Good; Bad	Weather condition affects visibility as well as the radar signal
Visibility	Adequate; Less than adequate	Visibility for visual detection
Size of encounter ship 1(2)	Big; Small	Size of the encounter ship influences the signal quality they generate on radar
Communication	Yes; No	Communication (through VHF) between the encountering ships about passing plan
Steering failure 1(2)	Yes; No	Whether there is a steering failure
Navigational complexity	High; Low	Complexity for navigation, depending on traffic intensity and geographical situation
Time of the day	Daytime; Night	Influences visibility

the red bars indicate the interval probabilities. Note that due to the limitation of space, Fig. 4 only shows half of the model presented in Fig. 2. One of two identical sets of nodes corresponding to each of the encountering ship was not shown in Fig. 4. The width of the red bar reflects the ambiguity of the experts' judgement. The wider the red bar, the more ambiguous the result would be. In the extreme cases of point probabilities for some nodes, the width of the red bar is zero, reflecting no ambiguity. For example, the interval probability for "Detection = Yes" is quite small, ranging from 0.966 to 0.999. However, for the node "Collision = Yes", the probability number is large, which is somewhere between 0.359 and 0.886, varying from "Medium likelihood" to "Likely to happen".

The wide range of the interval marginal probabilities could be partially explained by the use of the linguistic terms. The initial imprecision in conditional probabilities sequentially propagates to the posterior probabilities. Therefore, it might be advisable to carefully modify the terms in the future so as to allow more flexibility in probability specification. The ideal case should be to elicit the probability interval numbers directly where the experts can decide the broadness the interval for each judgment. However, allowing for flexibility would add to much complexity in application. In the first version elicitation tool, the experts were asked to specify the boundaries of conditional probability numbers by dragging two sliders. But the feedback from the experts is that they seem to be confused by too many numbers. It is always not easy to strike a balance. More effort should therefore be focused on this for future exploration.

4.2. Comparisons of results from different experts

In addition to the computation of the combined probabilities, separate calculations were also performed to compare the inputs provided by each of the eleven experts. We denoted "0" for the case using the combined interval conditional probability as input and "1", "2", … "10", "11" for the cases with inputs from expert "1", "2", … "10", "11". The marginal probabilities for six nodes ("Collision", "Maneuvering1", "Detection1", "Navigational system detection1", "Visual detection1" and "VTS detection") were chosen for comparison, as illustrated in Fig. 5.

The marginal probabilities of the same node from different experts in Fig. 6 show discrepancies, both in the width and ranges of the intervals, which further verifies the existence of epistemic uncertainty in the result of risk assessment. This research is the first in exploring the quantification of the epistemic uncertainty in maritime accidents risk modelling. To reduce

Table 4
Scales of linguistic terms with appropriate probability intervals.

(1) Mastrandrea et al. (2010)		(2) Ren et al. (2009)		(3) Antonucci et al. (2013b)	
Virtually certain	[99%, 100%]	Nearly certain	[95%, 100%]	Most likely	[90%, 100%]
Very likely	[90%, 99%]	Very likely	[70%, 95%]	Very likely	[80%, 90%]
Likely	[66%, 90%]	Quite likely	[55%, 70%]	Likely	[65%, 80%]
About as likely as not	[33%, 66%]	Even chance	[45%, 55%]	Fifty-fifty	[35%, 65%]
Unlikely	[10%, 33%]	Quite unlikely	[30%, 45%]	Unlikely	[20%, 35%]
Very unlikely	[1%, 10%]	Very unlikely	[5%, 30%]	Very unlikely	[10%, 20%]
Exceptionally unlikely	[0%, 1%]	Nearly impossible	[0%, 5%]	Most unlikely	[0%, 10%]

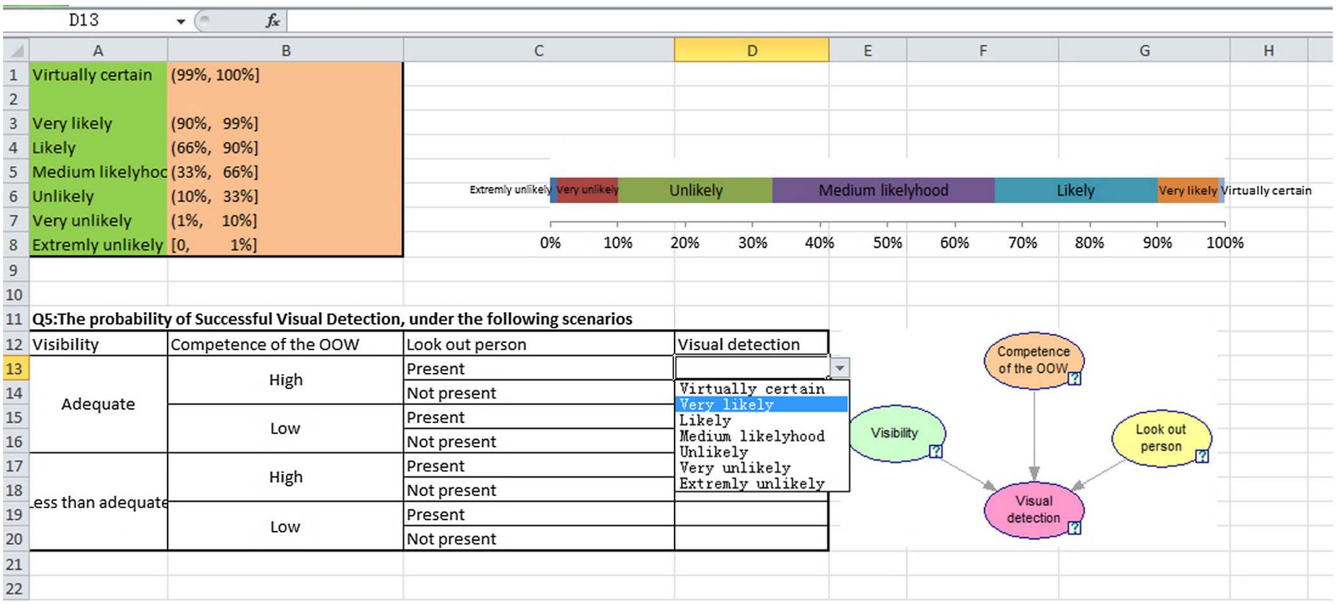

Fig. 3. Excerpt of the questionnaire for conditional probabilities elicitation.

Table 5
Prior probabilities for nodes without parent nodes.

Node	Prior probabilities	Description	References
Time of day	50% day; 50% night	–	–
Weather	49% Rain, 51% Clear	Mean Rain days, 178 days/365 days from 1891–2011 (121 yrs)	Singapore (2011)
VTS Presence	100% Present, 0% Absent	Mandatory Ship Reporting System in the Straits of Malacca and Singapore known as "STRAITREP	Segar Abdullah (2000)
Communication between ships	30% Yes; 70% No	Whether the vessel receives a communication call from the encountering vessel	Det Norske Veritas (2003)
Navigational complexity	80% High, 20% Low	General assumption considering the business of Singapore strait	–
Safety culture	50% Good; 50% Bad		
Size of encounter ship	98% Big, 2% Small	Average percentage of small fishing vessels, tug/tow/government vessel and others from 2009–2014	Rusli (2012)

Table 6
Prior interval probabilities for all nodes.

Nodes & states	Point	Interval	Nodes & states	Point	Interval
Pr(Collision = No)	0.406	[0.114, 0.641]	Pr(Maintenanceroutine1 = NotFollowed)	0.447	[0.383, 0.510]
Pr(Collision = Yes)	0.594	[0.359, 0.886]	Pr(Maintenanceroutine1 = Followed)	0.553	[0.490, 0.617]
Pr(VTSDetection = No)	0.062	[0.015, 0.109]	Pr(Resthours1 = LessThanAdequate)	0.305	[0.201, 0.420]
Pr(VTSDetection = Yes)	0.938	[0.891, 0.985]	Pr(Resthours1 = Adequate)	0.695	[0.580, 0.799]
Pr(Visibility = LessThanAdequate)	0.453	[0.376, 0.530]	Pr(Lookoutperson1 = No)	0.457	[0.336, 0.582]
Pr(Visibility = Adequate)	0.547	[0.470, 0.624]	Pr(Lookoutperson1 = Yes)	0.543	[0.418, 0.664]
Pr(Detection1 = No)	0.011	[0.001, 0.034]	Pr(Updatingroutine1 = NotFollowed)	0.430	[0.365, 0.495]
Pr(Detection1 = Yes)	0.989	[0.966, 0.999]	Pr(Updatingroutine1 = Followed)	0.570	[0.505, 0.635]
Pr(Maneuverplanning1 = Incorrect)	0.398	[0.215, 0.623]	Pr(Visualdetection1 = No)	0.340	[0.150, 0.566]
Pr(Maneuverplanning1 = Correct)	0.602	[0.377, 0.785]	Pr(Visualdetection1 = Yes)	0.660	[0.434, 0.850]
Pr(Maneuvering1 = Incorrect)	0.588	[0.413, 0.875]	Pr(Navigationalsystemdetection1 = No)	0.387	[0.252, 0.553]
Pr(Maneuvering1 = Correct)	0.412	[0.125, 0.587]	Pr(Navigationalsystemdetection1 = Yes)	0.613	[0.447, 0.748]
Pr(CompetenceoftheOOW1 = Low)	0.466	[0.273, 0.691]	Pr(Distractionlevel1 = Low)	0.509	[0.393, 0.658]
Pr(CompetenceoftheOOW1 = High)	0.534	[0.309, 0.727]	Pr(Distractionlevel1 = High)	0.491	[0.342, 0.607]
Pr(Steeringfailure1 = No)	0.614	[0.490, 0.732]	Pr(Bridgeresourcemanagement1 = Bad)	0.431	[0.357, 0.506]
Pr(Steeringfailure1 = Yes)	0.386	[0.268, 0.510]	Pr(Bridgeresourcemanagement1 = Good)	0.569	[0.494, 0.643]
Pr(Fatigue1 = No)	0.670	[0.520, 0.810]	Pr(Crewtraining1 = LessThanAdequate)	0.442	[0.370, 0.515]
Pr(Fatigue1 = Yes)	0.330	[0.190, 0.480]	Pr(Crewtraining1 = Adequate)	0.558	[0.485, 0.630]
Pr(Navigationalsystemsignal1 = Bad)	0.251	[0.106, 0.439]	Pr(Navigationalsystemsettings1 = Incorrect)	0.401	[0.199, 0.654]
Pr(Navigationalsystemsignal1 = Good)	0.749	[0.561, 0.894]	Pr(Navigationalsystemsettings1 = Correct)	0.599	[0.346, 0.801]

the epistemic uncertainty, careful attention should be given to all steps of risk assessment including the selection of experts, the education of experts on Bayesian Network as well as the elicitation process.

4.3. Backward inferences

Back forward inference was performed for the node "Collision". Changes in posterior probability of some nodes were shown in Fig. 6. Unlike the traditional Bayesian network method, it is not easy to draw conclusion on the exact change of posterior probabilities with the interval probabilities. Only differences in the upper and lower bounds could be observed implicitly. The biggest change in the bounds of the interval posterior probabilities was for the node "Maneuvering". Given the evidence "Collision = Yes", both the upper bound and lower

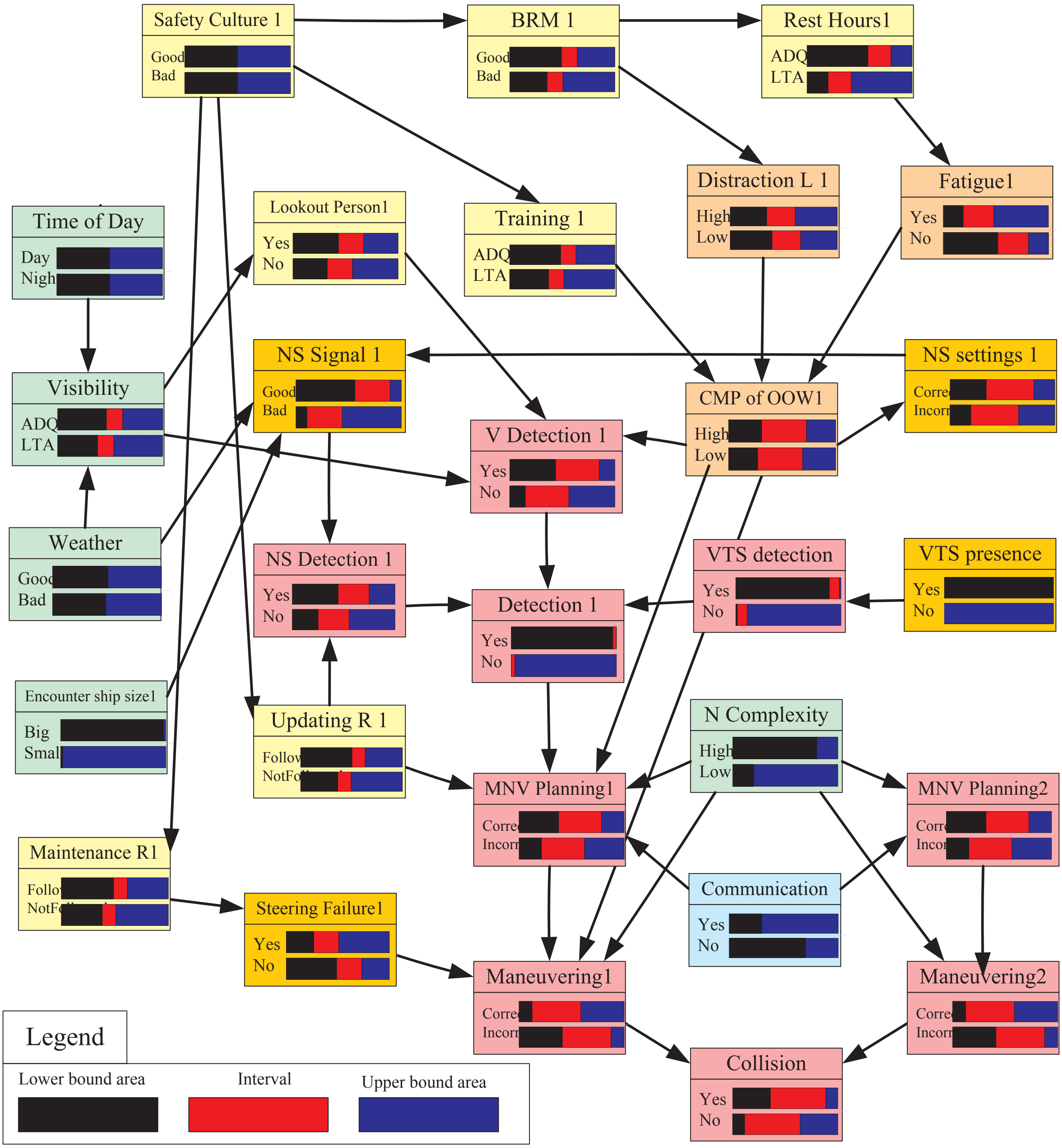

Fig. 4. Posterior interval probabilities for the collision model.

bounds for correct maneuvering have significantly decrease. On the other hand, there is a large increase in the bounds when the evidence was set as "Collision = No".

4.4. Evaluation of risk control option

The uncertainty in the result of risk assessment also propagates to the decision-making process and influences the ranking of potential risk control options. The following example is illustrated.

Measures recommended for the evaluation of risk control options (step four of the FSA) are Gross Cost of Averting a Fatality (GCAF) and Net Cost of Averting a Fatality (NCAF), which are defined as follows (Kontovas and Psaraftis, 2009; Psaraftis, 2012):

$$GCAF = \frac{\Delta C}{\Delta R} \tag{13}$$

$$NCAF = \frac{\Delta C - \Delta B}{\Delta R} \tag{14}$$

ΔC is the cost per ship for the implementation of the risk control options;
ΔB is the economic benefits resulting from the implementation, which include the saved cost for cargo damage, ship repairs, etc.;
ΔR is the risk reduction per ship.

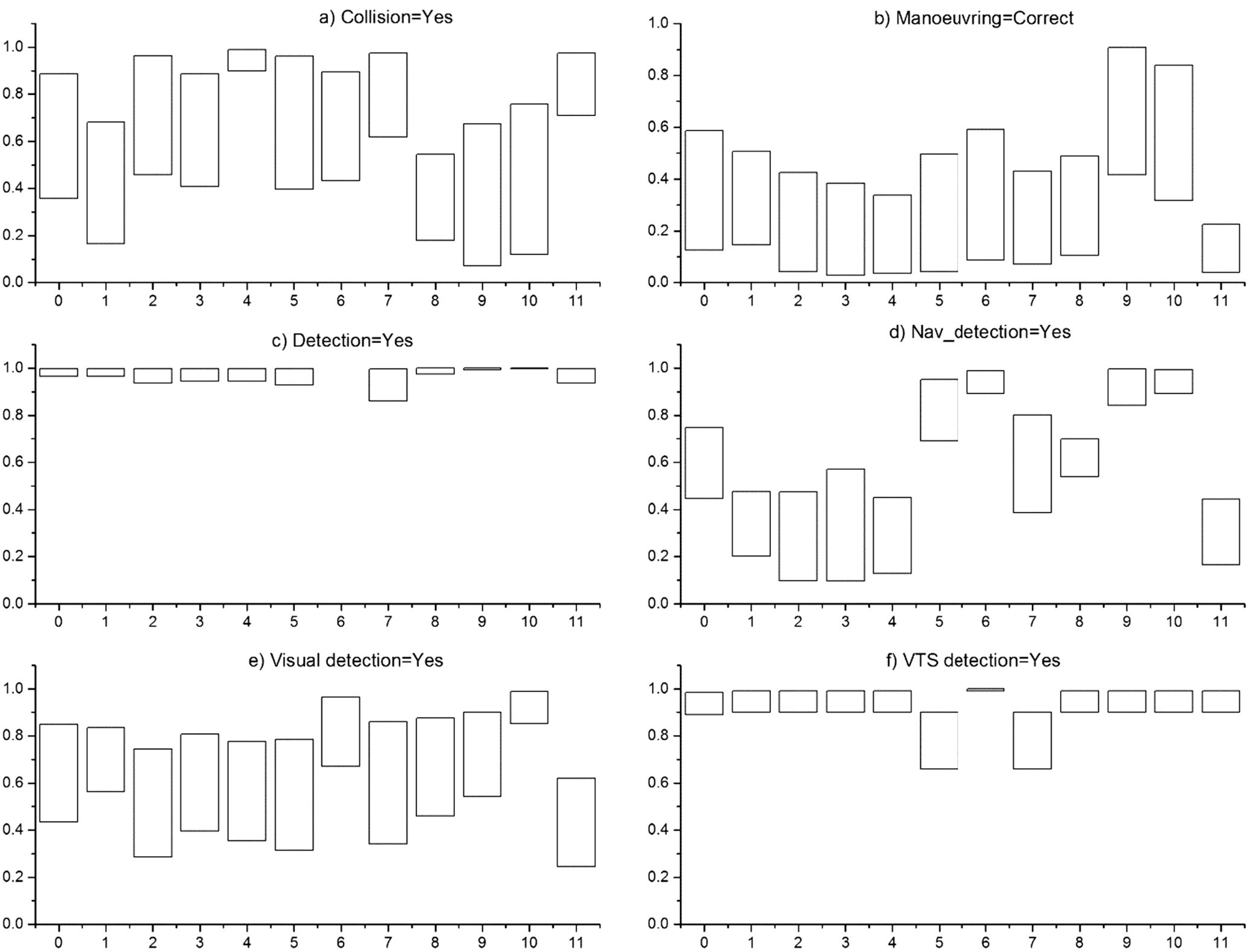

Fig. 5. Comparisons of prior probability intervals from all experts.

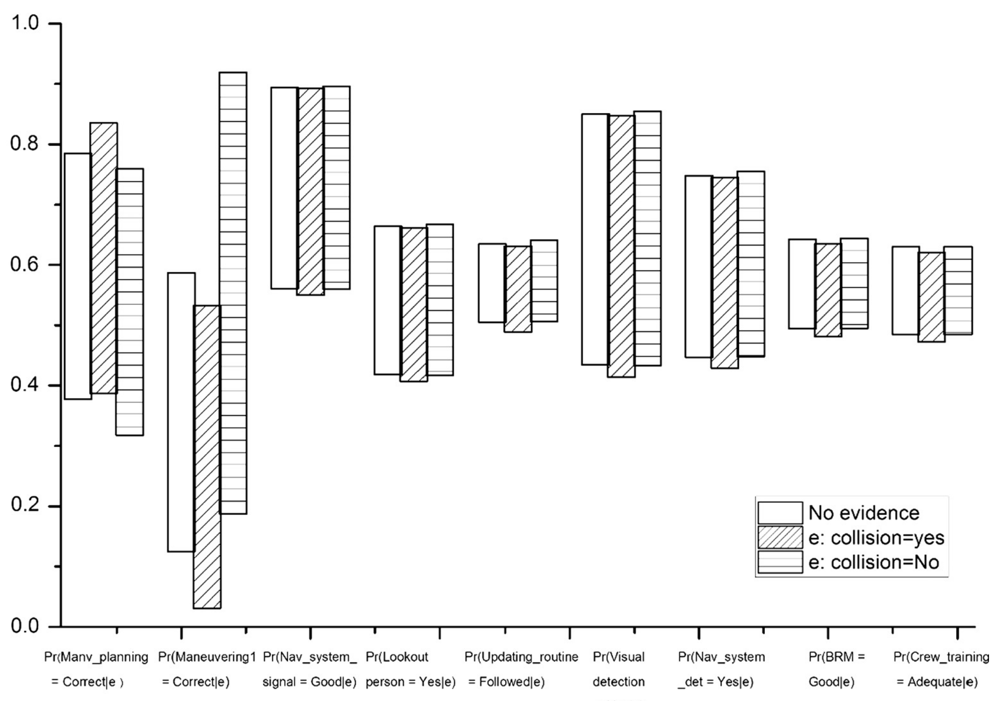

Fig. 6. Changes of posterior probability with observations of the node "collision"

Table 7
Cost expressions.

Linguistic terms	1	2	3	4	5	6	7
Very high	0	0	0	0	0	0.75	1
High	0	0	0	0	0.75	1	0.25
Moderately high	0	0	0	0.5	1	0.25	0
Average	0	0	0.5	1	0.5	0	0
Moderately low	0	0.25	1	0.5	0	0	0
Low	0.25	1	0.75	0	0	0	0
Very low	1	0.75	0	0	0	0	0

Table 8
Utility expressions.

Linguistic terms	1	2	3	4	5	6	7
Slightly preferred	0	0	0	0	0	0.75	1
Moderately preferred	0	0	0	0.5	1	0.25	0
Preferred	0	0.25	1	0.5	0	0	0
Greatly preferred	1	0.75	0	0	0	0	0

Table 9
Risk reduction for each RCO under interval and point Bayesian network.

RCOs	Lower bound	Upper bound	Point probability	Change of lower bound	Change of upper bound	Change of point probability
RCO1	0.273	0.839	0.529	0.086	0.047	0.064
RCO 2	0.274	0.838	0.531	0.085	0.048	0.063
RCO 3	0.313	0.865	0.544	0.046	0.021	0.049
RCO 4	0.271	0.835	0.518	0.088	0.051	0.076

Table 10
Cost evaluation.

RCOs	1	2	3	4	5	6	7
RCO1	0	0	0	0.6	1	0.5	0
RCO 2	0	0	0	0	0.7	1	0.4
RCO 3	0	0	0.55	1	0.45	0	0
RCO 4	0	0	0	0	0	0.6	1

Table 11
The RP values and rankings of RCO for different cases.

	RCO1	RCO2	RCO3	RCO4	Ranks
Case1	21.89	22.12	35.42	34.62	RCO1, RCO2, RCO4, RCO3
Case2	40.05	39.18	77.58	59.74	RCO2, RCO1, RCO4, RCO3
Case3	29.28	29.99	33.04	40.28	RCO1, RCO2, RCO3, RCO4

The dominant yardstick in all FSA studies that have been submitted to IMO so far for deciding a potential risk control option is the "\$3m criterion" (Kontovas and Psaraftis, 2009). Recent applications also include environmental consequences into consideration. The level of risk reduction ΔR could be obtained with the BN analysis but it is very difficult to obtain the exact cost and benefit values. The subjective method proposed by Wang et al. (2013) is adopted for analysis in this research since it does not require absolute cost and benefit values. The cost and benefit values can be defined with linguistic terms such as "High", "Low", etc., which could then be mapped onto the defined utility expressions. Definitions of the cost expression and the utility expression could be found in Tables 7 and 8, where the number 1–7 are 7 categories to which the linguistic variables could be mapped with a membership function/number. The categories don't have practical meaning but provide engineers with measures with which a linguistic variable can be modelled. For simplification, the ranks of risk control options in the following example will be based on GCAF only. Thus, only cost values need to be evaluated. The preference degree P of a risk control option could be calculated from the utility values (Wang et al., 2013). The final ranking of RCOs depends on the value of $RP_i = 1/(Pi \times \Delta R)$. The smaller the RP_i, the more cost effective RCO will be.

Four risk control options could be proposed based on the analysis of the model: (1) Simulator training for officers on difficult tasks such as passage planning and operations during malfunction of critical technical equipment; (2) Improving the bridge resource management for the bridge team; (3) Redundant propulsion or steering system; and (4) Improving safety culture. Changes in collision probabilities after implementing these risk control options were calculated with the traditional Bayesian Network method as well as the interval Bayesian Network method separately, as shown in Table 9. For risk reduction calculated with interval probabilities, the changes of the upper bound and lower bound were used as two special cases for comparison.

Suppose that the costs of these four RCOs (see Table 10) are "Moderately High", "High", "Average", "Very high" (Wang et al., 1996), their values are as follows:

The preference degree values P_i calculated with the best fit method for the RCO1 to RCO4 are 0.531218565, 0.531756404, 0.613794117, and 0.328221365 respectively. Combining with the risk reduction values from Table 9, i.e. "Change of Lower bound" (Case 1), "Change of Upper bound" (Case 2), and "Change of Point Probability" (Case 3), the RP values and rankings of RCO for different cases could be calculated as summarized in Table 11.

It could be seen from Table 11 that the rank of cost effectiveness values for different RCOs differs for the three cases. The example is just to simulate the impact of uncertainty in risk analysis result on decision making process. The detailed decision-making with the imprecise probabilities still remains a question for future research.

5. Conclusions

Risk prediction with BN is a popular methodology for modelling causal relationship, interdependence and easy updating. Experts' knowledge is the main source of data for BN construction and parameterization for many maritime accidents modelling due to the lack of empirical data. However, experts' involvement brings a high level of epistemic uncertainty. This paper explored the application of interval probabilities instead of tradition point probabilities in Bayesian Network to address this uncertainty. Linguistic terms defined with probability intervals were used for the probability elicitation process. It was found that experts feel more confident in providing their judgment with linguistic terms compared to point probability numbers. Inferences were made directly with the interval probabilities with the GL2U algorithm. The methodology was applied to the causation probability modelling of ship collision. The calculated posterior probabilities were also in the form of interval probabilities, representing imprecision in the results. The influences of uncertainty in risk assessment results on ranks of potential risk control options were discussed using a simple example. This is the first research which applies imprecise probabilities in BN-based maritime risk predictions. The methodology adds more values to BN and its practical applications to the modelling of maritime accidents.

Acknowledgements

We acknowledge the Interdisciplinary Graduate School of Nanyang Technological University (NTU) for providing the research scholarship for the first author as a PhD candidate. We appreciate the financial support and research facility from DHI-NTU Centre, under the umbrella of Nanyang Environment and Water Research Institute (NEWRI) of NTU. We would like to thank those experts who were interviewed and have contributed their precious time to discussions and answering the questionnaires. Last but not the least, we would also like to express our sincere gratitude to our reviewers for their valuable comments and suggestions which are of great importance to improving the quality of the paper.

Appendix A

See Table A1.

Table A1
Parent and child nodes for all nodes in the model.

Node	Parent node	Child node
Collision	Maneuvering 1(2)	
Maneuvering 1(2)	Steering failure 1(2); Competence of the OOW 1(2); Maneuver planning 1(2); Navigational complexity	Collision
Maneuver planning 1(2)	Detection (2); Competence of the OOW 1(2); Navigational complexity; Communication between the ships	Maneuvering 1(2)
Updating routine 1(2)	Safety culture 1(2)	Navigational system detection 1(2)
Navigational system signal 1(2)	Navigational System settings 1(2); Size of encounter ship 1(2); Weather	Navigational system detection 1(2)
Navigational System settings 1(2)	Competence of the OOW 1(2)	Navigational system signal 1(2)
Navigational system detection 1(2)	Navigational system signal 1(2); Updating routine 1(2)	Detection 1(2)
Look out person 1(2)	Bridge Resource Management BRM 1(2); Visibility	Visual detection 1(2)
Visual detection 1(2)	Competence of the OOW 1(2); Look out person 1(2); Visibility	Detection 1(2)
VTS Presence		VTS Detection
VTS Detection	VTS Presence	Detection 1(2)
Detection 1(2)	Navigational system detection 1(2); Visual detection 1(2); VTS Detection	Maneuver planning 1(2)
Rest hours 1(2)	Bridge Resource Management BRM 1(2)	Fatigue 1(2)
Fatigue 1(2)	Rest hours 1(2)	Competence of the OOW 1(2)
Distraction level 1(2)	Bridge Resource Management BRM 1(2)	Competence of the OOW 1(2)
Crew training 1(2)	Safety culture 1(2)	Competence of the OOW 1(2)
Competence of the OOW 1(2)	Fatigue 1(2); Distraction level 1(2); Crew training 1(2)	Navigational System settings 1(2); Visual detection 1(2); Maneuver planning 1(2); Maneuvering 1(2)
Safety culture 1(2)		Crew training 1(2); Bridge Resource Management BRM 1(2); Updating routine 1(2); Maintenance routine 1(2)
Maintenance routine 1(2)	Safety culture 1(2)	Steering failure 1(2)
Bridge Resource Management BRM 1(2)	Safety culture 1(2)	Rest hours 1(2); Distraction level 1(2)
Weather		Visibility; Navigational system signal 1(2)
Visibility	Time of the day; Weather	Look out person 1(2); Visual detection 1(2)
Size of encounter ship 1(2)		Navigational system signal 1(2)
Communication		Maneuver planning 1(2)
Steering failure 1(2)	Maintenance routine 1(2)	Maneuvering 1(2)
Navigational complexity		Maneuver planning 1(2); Maneuvering 1(2)
Time of the day		Visibility

Appendix B. Combined interval probabilities from all experts

See Tables B1–B19.

Table B1
The combined interval conditional probabilities for the node “Collision”

Manoeuvre of ship 1	Manoeuvre of ship 2	Collision = Yes	
		Lower bound	Upper bound
Correct	Correct	0.018	0.079
	Incorrect	0.434	0.714
Incorrect	Correct	0.434	0.714
	Incorrect	0.832	0.952

Table B2
The combined interval conditional probabilities for the node “Maneuvering”

Manoeuvre planning	Steering failure	Competence of the OOW	Navigational complexity	Maneuvering = Correct	
				Lower bound	Upper bound
Correct	Yes	High	High	0.394	0.664
			Low	0.595	0.762
		Low	High	0.125	0.233
			Low	0.192	0.384
	No	High	High	0.839	0.933
			Low	0.921	0.947
		Low	High	0.212	0.438
			Low	0.399	0.663

Table B3
The combined interval conditional probabilities for the node "Maneuver planning"

Detection	Navigational complexity	Competence of the OOW	Communication between ships	Manoeuvre planning = Correct	
				Lower bound	Upper bound
Yes	High	High	Yes	0.903	0.986
			No	0.714	0.871
		Low	Yes	0.357	0.635
			No	0.098	0.281
	Low	High	Yes	0.968	0.998
			No	0.890	0.974
		Low	Yes	0.480	0.731
			No	0.264	0.509

Table B4
The combined interval conditional probabilities for the node "Navigational system detection"

Updating routine	Navigational system signal	Navigational system detection = Yes	
		Lower bound	Upper bound
Followed	Good	0.937	0.994
	Bad	0.445	0.624
Not followed	Good	0.275	0.420
	Bad	0.032	0.112

Table B5
The combined interval conditional probabilities for the node "Visual detection"

Visibility	Competence of the OOW	Look out person	Visual detection = Yes	
			Lower bound	Upper bound
Adequate	High	Present	0.911	0.990
		Not present	0.761	0.927
	Low	Present	0.760	0.936
		Not present	0.327	0.615
Less than adequate	High	Present	0.719	0.901
		Not present	0.357	0.623
	Low	Present	0.327	0.598
		Not present	0.073	0.229

Table B6
The combined interval conditional probabilities for the node "VTS detection"

VTS Presence	VTS detection = Yes	
	Lower bound	Upper bound
Present	0.891	0.985
Not present	0.019	0.068

Table B7
The combined interval conditional probabilities for the node "Rest hours"

BRM	Rest hours = Adequate	
	Lower bound	Upper bound
Good	0.922	0.991
Bad	0.246	0.454

Table B8
The combined interval conditional probabilities for the node "Fatigue"

Rest hours	Fatigue = Yes	
	Lower bound	Upper bound
Adequate	0.027	0.133
Less than adequate	0.837	0.961

Table B9
The combined interval conditional probabilities for the node "Distraction"

BRM	Level of distraction = High	
	Lower bound	Upper bound
Good	0.111	0.281
Bad	0.759	0.925

Table B10
The combined interval conditional probabilities for "Updating routine"

Safety culture	Updating routine = Followed	
	Lower bound	Upper bound
Good	0.914	0.987
Bad	0.097	0.284

Table B11
The combined interval conditional probabilities for "OOW Competence"

Crew training	Fatigue	Distraction level	Competent OOW = High	
			Lower bound	Upper bound
Adequate	Yes	High	0.108	0.310
		Low	0.220	0.471
	No	High	0.676	0.865
		Low	0.911	0.987
Less than adequate	Yes	High	0.006	0.063
		Low	0.099	0.277
	No	High	0.165	0.416
		Low	0.408	0.699

Table B12
The combined interval conditional probabilities for "Maintenance routine"

Safety culture	Maintenance routine = Followed	
	Lower bound	Upper bound
Good	0.895	0.985
Bad	0.086	0.248

Table B13
The combined interval conditional probabilities for the node "BRM"

Safety culture	BRM = Good	
	Lower bound	Upper bound
Good	0.877	0.978
Bad	0.112	0.308

Table B14
The combined interval conditional probabilities for the node "Training"

Safety culture	Training = Adequate	
	Lower bound	Upper bound
Good	0.872	0.978
Bad	0.098	0.281

Table B15
The combined interval conditional probabilities for the node "Look out"

Visibility	BRM	Look out person = Present	
		Lower bound	Upper bound
Adequate	Good	0.580	0.689
	Bad	0.022	0.122
Less than adequate	Good	0.953	0.996
	Bad	0.286	0.509

Table B16
The combined interval conditional probabilities for the node "Visibility"

Time of the day	Weather	Visibility = Good	
		Lower bound	Upper bound
Daytime	Good	0.922	0.991
	Bad	0.193	0.451
Night	Good	0.711	0.906
	Bad	0.027	0.119

Table B17
The combined interval conditional probabilities for "Steering failure"

Maintenance routine	Steering failure = Yes	
	Lower bound	Upper bound
Followed	0.034	0.151
Not followed	0.644	0.855

Table B18
The combined interval conditional probabilities for "Navigational system signal"

Weather	Navigational system settings	Size of encounter ship	Navigational system signal = Good	
			Lower bound	Upper bound
Good	Correct	Big ships	0.937	0.994
		Small vessels	0.692	0.911
	Incorrect	Big ships	0.475	0.646
		Small vessels	0.109	0.275
Bad	Correct	Big ships	0.842	0.961
		Small vessels	0.340	0.649
	Incorrect	Big ships	0.316	0.511
		Small vessels	0.004	0.037

Table B19
The combined interval conditional probabilities for the node "Navigational system settings"

Competence of OOW	Navigational system settings = Correct	
	Lower bound	Upper bound
High	0.911	0.990
Low	0.092	0.299

References

Akhtar, M.J., Utne, I.B., 2014. Human fatigue's effect on the risk of maritime groundings – a Bayesian network modeling approach. Saf. Sci. 62, 427–440.

Antonucci, A., 2008. Imprecise Probabilistic Graphical Models: Equivalent Representations, Inference Algorithms and Applications. University of Lugano.

Antonucci, A., De Campos, C.P., Huber, D., Zaffalon, M., 2013a. Approximating credal network inferences by linear programming. In: Symbolic and Quantitative Approaches to Reasoning with Uncertainty. Springer, pp. 13–24.

Antonucci, A., Huber, D., Zaffalon, M., Luginbuhl, P., Chapman, I., Ladouceur, R., 2013b. Credo: a military decision-support system based on credal networks. In: Proceedings of the 16th International Conference on Information Fusion (FUSION 2013), Istanbul, Turkey, pp. 1942–1949.

Antonucci, A., Sun, Y., De Campos, C.P., Zaffalon, M., 2010. Generalized loopy 2U: a new algorithm for approximate inference in credal networks. Int. J. Approximate Reasoning 51 (5), 474–484.

Antonucci, A., Zaffalon, M., 2008. Decision-theoretic specification of credal networks: a unified language for uncertain modeling with sets of Bayesian networks. Int. J. Approximate Reasoning 49 (2), 345–361.

Banda, O.a.V., Goerlandt, F., Kuzmin, V., Kujala, P., Montewka, J., 2016. Risk management model of winter navigation operations. Mar. Pollut. Bull. 108 (1), 242–262.

Bayesian Fusion, L., 2017. < https://www.Bayesfusion.Com/genie-modeler > .

Corani, G., Antonucci, A., Zaffalon, M., 2012. Bayesian Networks with Imprecise Probabilities: Theory and Application to Classification. Data Mining: Foundations and Intelligent Paradigms. Springer, pp. 49–93.

Cozman, F.G., 2000. Credal networks. Artif. Intell. 120 (2), 199–233.

De Campos, L.M., Huete, J.F., Moral, S., 1994. Uncertainty management using probability intervals. In: Advances in Intelligent Computing—ipmu'94. Springer, pp. 190–199.

Det Norske Veritas, 2003. Formal safety assessment-large passenger ships, annex ii: risk assessment-large passenger ships navigation.

Fagiuoli, E., Zaffalon, M., 1998. 2U: an exact interval propagation algorithm for polytrees with binary variables. Artif. Intell. 106 (1), 77–107.

Fallet, G., Duval, C., Simon, C., Weber, P., Iung, B., 2011. Expert judgments collecting and modeling: application to the integrated risks analysis (IRA) methodology. In: Proceedings of the 3rd International Workshop on Dependable Control of Discrete Systems (DCDS), Saarbrucken, Germany, pp. 72–77.

Friis-Hansen, P., Simonsen, B.C., 2002. Gracat: software for grounding and collision risk analysis. Mar. Struct. 15 (4), 383–401.

Fujii, Y., Oshima, R., Yamanouchi, H., Mizuki, N., 1974. Some factors affecting the frequency of accidents in marine traffic: i-the diameter of evasion for crossing encounters, ii-the probability of stranding, iii-the effect of darkness of the probability of collision and stranding. J. Navigat. 27 (2), 239–247.

Ge, Y., Zhou, L., Cui, S., Yang, M., 2013. Extending genie for building graphical decision-theoretic models. In: Proceedings of the 10th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Shenyang, China, pp. 588–593.

Goerlandt, F., Kujala, P., 2011. Traffic simulation based ship collision probability modeling. Reliab. Eng. Syst. Saf. 96 (1), 91–107.

Goerlandt, F., Montewka, J., 2015a. A framework for risk analysis of maritime transportation systems: a case study for oil spill from tankers in a ship–ship collision. Saf. Sci. 76, 42–66.

Goerlandt, F., Montewka, J., 2015b. Maritime transportation risk analysis: review and analysis in light of some foundational issues. Reliab. Eng. Syst. Saf. 138, 115–134.

Grabowski, M., 2000. Risk modeling in distributed, large-scale systems. Citeseer.

Guo, P., Tanaka, H., 2010. Decision making with interval probabilities. Eur. J. Oper. Res. 203 (2), 444–454.

Hall, J., Twyman, C., Kay, A., 2005. Influence diagrams for representing uncertainty in climate-related propositions. Clim. Change 69 (2–3), 343–365.

Helle, I., Lecklin, T., Jolma, A., Kuikka, S., 2011. Modeling the effectiveness of oil combating from an ecological perspective – a Bayesian network for the Gulf of Finland; the Baltic Sea. J. Hazard. Mater. 185 (1), 182–192.

Hu, X., Luo, H., Fu, C., 2012. Probability elicitation in influence diagram modeling by using interval probability. Int. J. Intellig. Sci. 2 (04), 89–95.

Ide, J.S., Cozman, F.G., 2004. Ipe and l2u: approximate algorithms for credal networks. In: Onaindia, E., Staab, S. (Eds.), Stairs 2004: Proceedings of the Second Starting AI Researchers' Symposium. IOS Press, pp. 118–127.

Ide, J.S., Cozman, F.G., 2008. Approximate algorithms for credal networks with binary variables. Int. J. Approximate Reasoning 48 (1), 275–296.

IMO, 2006. Formal safety assessment. Consideration on utilization of Bayesian network at step 3 of FSA. Maritime Safety Committee 81st Session, MSC 81/18/1.

Kjrćulff, U.B., Madsen, A.L., 2013. Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis, second ed. Springer, New York, USA.

Kontovas, C.A., Psaraftis, H.N., 2009. Formal safety assessment: a critical review. Mar. Technol. 46 (1), 45–59.

Kuehmayer, J., 2008. Marine accident and casualty investigation boards. Reports, data bases, statistics, trends, lessons learned.

Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecol. Lett. 13 (7), 900–914.

Kujala, P., Hänninen, M., Arola, T., Ylitalo, J., 2009. Analysis of the marine traffic safety in the Gulf of Finland. Reliab. Eng. Syst. Saf. 94 (8), 1349–1357.

Lehikoinen, A., HäNninen, M., StorgåRd, J., Luoma, E., MäNtyniemi, S., Kuikka, S., 2015. A Bayesian network for assessing the collision induced risk of an oil accident in the Gulf of Finland. Environ. Sci. Technol. 49 (9), 5301–5309.

Leva, M., Friis-Hansen, P., Ravn, E.S., Lepsoe, A., 2006. Safedor: a practical approach to model the action of an officer of the watch in collision scenarios. In: Proceedings of the European Safety and Reliability Conference; ESREL 2006, Estoril, Portugal, pp. 2795–2804.

Li, S., Meng, Q., Qu, X., 2012. An overview of maritime waterway quantitative risk assessment models. Risk Anal. 32 (3), 496–512.

Liu, J., Yang, J., Wang, J., Sii, H., 2003. Review of uncertainty reasoning approaches as guidance for maritime and offshore safety-based assessment. J. UK Saf. Reliab. Soc. 23 (1), 63–80.

Liu, W.-Y., Yue, K., 2011. Bayesian network with interval probability parameters. Int. J. Artif. Intell. Tools 20 (05), 911–939.

Mastrandrea, M.D., Field, C.B., Stocker, T.F., Edenhofer, O., Ebi, K.L., Frame, D.J., Held, H., Kriegler, E., Mach, K.J., Matschoss, P.R., 2010. Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. In: Proceedings of the Intergovernmental Panel on Climate Change (IPCC), Jasper Ridge, CA, USA.

Mazaheri, A., 2009. Probabilistic modeling of ship grounding. Helsinki Univ. Technol. 48–56.

Merrick, J.R., Van Dorp, J.R., Dinesh, V., 2005. Assessing uncertainty in simulation-based maritime risk assessment. Risk Anal. 25 (3), 731–743.

Merz, B., Thieken, A.H., 2005. Separating natural and epistemic uncertainty in flood frequency analysis. J. Hydrol. 309 (1), 114–132.

Montewka, J., Ehlers, S., Goerlandt, F., Hinz, T., Tabri, K., Kujala, P., 2014. A framework for risk assessment for maritime transportation systems—a case study for open sea collisions involving ropax vessels. Reliab. Eng. Syst. Saf. 124, 142–157.

Montewka, J., Goerlandt, F., Innes-Jones, G., Owen, D., Hifi, Y., Puisa, R., 2017. Enhancing human performance in ship operations by modifying global design factors at the design stage. Reliab. Eng. Syst. Saf. 159, 283–300.

Neil, M., Fenton, N., Nielson, L., 2000. Building large-scale Bayesian networks. Knowl. Eng. Rev. 15 (3), 257–284.

Pedersen, P.T., 2010. Review and application of ship collision and grounding analysis procedures. Mar. Struct. 23 (3), 241–262.

Piatti, A., Antonucci, A., Zaffalon, M., 2010. Building knowledge-based systems by credal networks: a tutorial. Nova Sci.

Psaraftis, H.N., 2012. Formal safety assessment: an updated review. J. Mar. Sci. Technol. 17 (3), 390–402.

Ren, J., Jenkinson, I., Wang, J., Xu, D., Yang, J., 2008. A methodology to model causal relationships on offshore safety assessment focusing on human and organizational factors. J. Saf. Res. 39 (1), 87–100.

Ren, J., Jenkinson, I., Wang, J., Xu, D., Yang, J., 2009. An offshore risk analysis method using fuzzy Bayesian network. J. Offshore Mech. Arct. Eng. 131 (4), 12.

Ren, J., Wang, J., Jenkinson, I., Xu, D., Yang, J., 2007. A Bayesian network approach for offshore risk analysis through linguistic variables. Chin. Ocean Eng. 21 (3), 371–390.

Renooij, S., 2001. Probability elicitation for belief networks: issues to consider. Knowl. Eng. Rev. 16 (03), 255–269.

Rusli, M.H.M., 2012. Navigational hazards in international maritime chokepoints: a study of the straits of Malacca and Singapore. J. Int. Stud. 8 (47), 1–35.

Segar Abdullah, M., 2000. VTS [vessel traffic services] in the Singapore strait: an investigation into mandatory traffic control.

Singapore, N.E.A., 2011. Weather statistics.

Sormunen, O.-V.E., Goerlandt, F., Häkkinen, J., Posti, A., Hänninen, M., Montewka, J., Ståhlberg, K., Kujala, P., 2014. Uncertainty in maritime risk analysis: extended case study on chemical tanker collisions. Proc. Inst. Mech. Eng., Part M: J. Eng. Marit. Environ. 229 (3), 303–320.

Swiler, L.P., Paez, T.L., Mayes, R.L., 2009. Epistemic uncertainty quantification tutorial. In: Proceedings of the 27th International Modal Analysis Conference.

Tessem, B., 1992. Interval probability propagation. Int. J. Approximate Reasoning 7 (3), 95–120.

Wang, J., Yang, J., Sen, P., 1996. Multi-person and multi-attribute design evaluations using evidential reasoning based on subjective safety and cost analyses. Reliab. Eng. Syst. Saf. 52 (2), 113–128.

Wang, Y.F., Roohi, S.F., Hu, X.M., Xie, M., 2010. A new methodology to integrate human factors analysis and classification system with Bayesian network. In: Proceedings of the International Conference on Industrial Engineering and Engineering Management (IEEM), Macao, China, pp. 1776–1780.

Wang, Y.F., Xie, M., Chin, K.-S., Fu, X.J., 2013. Accident analysis model based on Bayesian network and evidential reasoning approach. J. Loss Prev. Process Ind. 26 (1), 10–21.

Weichselberger, K., 2000. The theory of interval-probability as a unifying concept for uncertainty. Int. J. Approximate Reasoning 24 (2), 149–170.

Zhang, D., Yan, X., Yang, Z., Wall, A., Wang, J., 2013. Incorporation of formal safety assessment and bayesian network in navigational risk estimation of the Yangtze River. Reliab. Eng. Syst. Saf. 118, 93–105.

Zhang, G., Thai, V.V., 2016. Expert elicitation and bayesian network modeling for shipping accidents: a literature review. Saf. Sci. 87, 53–62.